SQT

NASA

164000-10-P

Sixth Type II Quarterly Status
and Technical Report

STUDY ON SPECTRAL/RADIOMETRIC CHARACTERISTICS OF THE THEMATIC MAPPER FOR LAND USE APPLICATIONS

21 December 1983 — 20 March 1984

WILLIAM A. MALILA

APRIL 1984

(E84-10130) STUDY ON SPECTRAL/RADIOMETRIC CHARACTERISTICS OF THE THEMATIC MAPPER FOR LAND USE APPLICATIONS Quarterly Status Technical Progress Report, 21 Dec. 1983 - 20 Mar. 1984 (Environmental Research Inst. of N84-25140 Unclas G3/43 00130

Contract NAS5-27346
NASA Goddard Space Flight Center
Greenbelt Road
Greenbelt, Maryland 20771

ENVIRONMENTAL
RESEARCH INSTITUTE OF MICHIGAN
BOX 8618 • ANN ARBOR • MICHIGAN 48107

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	164000-10-P
2. Government Accession No.	
3. Recipient's Catalog No.	
4. Title and Subtitle	Study on Spectral/Radiometric Characteristics of the Thematic Mapper for Land Use Applications
5. Report Date	April 1984
6. Performing Organization Code	
7. Author(s)	William A. Malila
8. Performing Organization Report No.	164000-10-P
9. Performing Organization Name and Address	Environmental Research Institute of Michigan P.O. Box 8618 Ann Arbor, MI 48107
10. Work Unit No.	
11. Contract or Grant No.	NAS5-27346
12. Sponsoring Agency Name and Address	National Aeronautics and Space Administration Goddard Space Flight Center Greenbelt, MD 20771
13. Type of Report and Period Covered	Quarterly Status and Technical Progress 21 Dec 1983 - 20 Mar 1984
14. Sponsoring Agency Code	
15. Supplementary Notes	Mr. Harold Oseroff (Code 902) is serving as Technical Officer and Mssrs. Brian Markham and James Irons (Code 923) are serving as Science Representatives for NASA/GSFC.

16. Abstract

An information-theoretic measure of multispectral information content is developed and applied to a simultaneous Landsat TM and MSS data set. The entropy-based function measures the dispersion and concentration of signal values in various data spaces, irrespective of specific class memberships. It is used to compare the information content of TM vs. MSS and of various subsets of TM and MSS bands, as well as Tasseled-Cap transformations of the band values. Differences exist between the information-measure results and results using variance-based measures. Comparisons of system-design information capacities and data-space volumes also are made. The results and observations presented are considered to be preliminary in nature since only one real and one simulated data set were analyzed.

17. Key Words	Landsat 4 Thematic Mapper Multispectral Scanner Information Content Entropy
18. Distribution Statement	Initial distribution is listed at the end of this document.
19. Security Classif. (of this report)	UNCLASSIFIED
20. Security Classif. (of this page)	UNCLASSIFIED
21. No. of Pages	25 + iii
22. Price	

ERIM INFRARED AND OPTICS DIVISION

Report No. 164000-10-P

Sixth
Type II Quarterly Status
and Technical Progress Report
21 December 1983 - 20 March 1984

for

Study on Spectral/Radiometric Characteristics
of the Thematic Mapper for Land Use Applications

under

Contract NAS5-27346

with

NASA Goddard Space Flight Center
Greenbelt Road
Greenbelt, Maryland 20771

Submitted by

Environmental Research Institute of Michigan
P.O. Box 8618
Ann Arbor, Michigan 48107

Prepared by: William A. Malila
William A. Malila
Principal Investigator

Approved by: Robert Horvath
Robert Horvath
Manager, Information Processing
Department

April 1984

CONTENTS

Sixth Quarterly Report

STUDY ON SPECTRAL/RADIOMETRIC CHARACTERISTICS OF THE THEMATIC MAPPER FOR LAND USE APPLICATIONS

1. OBJECTIVE

The objective of this investigation is to quantify the performance of the TM as manifested by the quality of its image data, in order to suggest improvements in data production and to assess the effects of the data quality on its utility for land resources applications. Three categories of this analysis are: a) radiometric effects, b) spatial effects and c) geometric effects, with emphasis on radiometric effects.

2. TASKS

Four tasks have been established to address the above objective. The first three are to study radiometric performance, spatial performance and geometric performance, respectively, while the fourth is to study spectral characteristics. In keeping with the identified objective, the radiometric performance study is the major task.

3. STATUS AND TECHNICAL PROGRESS

During this sixth quarterly reporting period, efforts were concentrated on developing a measure of the information content of multispectral data, such as Thematic Mapper (TM) and Multispectral Scanner (MSS) data, and then comparing results obtained upon applying the measure to simultaneous data from TM and MSS.

3.1 PROBLEMS

None.

3.2 ACCOMPLISHMENTS

An information-theoretic measure of multispectral information content was developed and applied to simultaneous Landsat TM and MSS data sets and preliminary observations and comparisons were made.

3.2.1 OBJECTIVES

With multispectral data sets from remote sensing systems, questions arise as to the relative merits of individual and groups of spectral bands and transformed spectral variables. Classification-based measures

are frequently used for such comparisons, as are variance-based measures. The objectives of the work reported here were to develop a class-independent and non-parametric measure and to apply it to Landsat TM and MSS data sets; the measure developed is based on information theoretic principles.

3.2.2 APPROACH

A communications-theory approach is taken to analyze the dispersion and concentration of signal values in various data spaces, irrespective of any specific class memberships. Entropy, as defined by Shannon, is used as the basic measure of information. The process of selecting a subset of bands is viewed as the transmission of data through a lossy communication channel, and the mutual information between the input and output is the measure of information transfer, i.e., the information represented by the subset.

The new measure was applied to Landsat Multispectral Scanner (MSS) and six-band Thematic Mapper (TM) data of two types. These are simulated data values derived from field-measured reflectance spectra of agricultural crops and soils and an atmospheric model, and actual Landsat-4 MSS and TM data acquired simultaneously from an agricultural scene in North Carolina. These data were used in a prior comparison we made of the spatial and spectral characteristics of Landsat TM and MSS data [1,2].

Several different comparisons of information content are made. These include comparison of TM and MSS system-design information capacities, comparison of the data-space volumes spanned by the agricultural data in the spaces defined by original bands and by transformed spectral (Tasseled-Cap) variables, comparison of the agricultural information content of original bands to that of transformed variables, and comparison of the agricultural information content of TM to that of MSS.

3.2.3 INFORMATION MEASURE DERIVATION

3.2.3.1 <u>Basic Concepts</u>. Shannon defined self information, $I(x_i)$, as a measure of the information associated with knowing the occurrence of a signal state x_i which occurs with probability $P(x_i)$:

$$I(x_i) = \log_2 \left(\frac{1}{P(x_i)}\right) = -\log_2 P(x_i) \qquad \text{(bits)} \qquad (1)$$

The more rare the event, the greater is one's uncertainty about when it will occur and, consequently, the greater is the information conveyed

when it is observed. Entropy, given the symbol H, is the value of self information when averaged over all N possible states of x:

$$H(x) = \sum_{i=1}^{N} P(x_i) \log_2 \frac{1}{P(x_i)} \qquad (2)$$

With two variables, the use of joint and conditional probabilities is necessary:

$$H(x,y) = H(x) + H(y|x) \qquad (3a)$$

or

$$H(x,y) = H(y) + H(x|y) \qquad (3b)$$

since

$$P(x,y) = P(x)P(y|x) \qquad (4a)$$

or

$$P(x,y) = P(y)P(x|y) \qquad (4b)$$

In computing the conditional entropy, the weighting assigned to each information term is the joint probability of the states involved, i.e., for example,

$$H(x|y) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} P(x_i,y_j) \log_2 \frac{1}{P(x_i|y_j)} \qquad (5)$$

If we consider x to be the input to a communication channel and y to be the output, we can define the mutual information transferred between them, i.e., $I_M(x;y)$, as

$$I_M(x;y) = H(x) - H(x|y) \qquad (6)$$

In words, the mutual information exchanged is the difference between H(x), the information content of the input, and H(x|y), the uncertainty about x when we are given the output y. When the total information is transferred, $H(x|y) = 0$ and $I_M(x;y) = H(x)$. At the other extreme, when y does not contain any information relatable to x, $H(x|y) = H(x)$ and therefore $I_M(x;y) = 0$, i.e., the mutual information is zero.

A convenient measure of channel (signal transformation) efficiency is the relative entropy or the ratio of mutual information to the total information of the input:

$$M_r = \frac{I_M(x;y)}{H(x)} \qquad (7a)$$

$$M_r = \frac{H(x) - H(x|y)}{H(x)} = 1 - \frac{H(x|y)}{H(x)} \qquad (7b)$$

3.2.3.2 Multispectral Extension. The above concepts can be extended to multispectral variables by letting the variables x and y become multidimensional vectors X and Y, with $X = (X_1, X_2, \ldots, X_{N_x})$ and $Y = (Y_1, Y_2, \ldots, Y_{N_y})$. Usually, $N_y \leq N_x$. The transformation achieved by the communication channel is used here in a general sense, to represent both simple selections of spectral band subsets and more complex transformations, such as the Tasseled-Cap Transformation.

The entropy of the input {X} becomes a function of the frequency with which individual signal-space cells or states are populated. Since the Thematic Mapper has six reflective bands, the equations are presented here in terms of six variables, although they should be adaptable to any number. The total information is:

$$H(X) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sum_{k=1}^{N_3} \sum_{l=1}^{N_4} \sum_{m=1}^{N_5} \sum_{n=1}^{N_6} P(X_{1i}, X_{2j}, X_{3k}, X_{4l}, X_{5m}, X_{6n})$$

$$\cdot \log_2 \frac{1}{P(X_{1i}, X_{2j}, X_{3k}, X_{4l}, X_{5m}, X_{6n})} \tag{8a}$$

where $P(X_{1i}, X_{2j}, X_{3k}, X_{4l}, X_{5m}, X_{6n})$ is the frequency with which state X(ijklmn) is populated,

and N_p is the number of populated levels of variable X_p.

To shorten subsequent equations, abbreviated notation will at times be used, e.g.,

$$H(X) = \sum_{ijklmn} \cdots \sum P_X(ijklmn) \log_2 \frac{1}{P_X(ijklmn)} \tag{8b}$$

The total number of possible states or cells,

$$N_{cells} = N_1 N_2 N_3 N_4 N_5 N_6$$

can be very large, but the vast majority will not be populated. From a calculational standpoint,

$$P_X(ijklmn) = \frac{C_{ijklmn}}{\sum \cdots \sum_{ijklmn} C_{ijklmn}} = \frac{C_{ijklmn}}{N_{obs}} \tag{9}$$

where C_{ijklmn} is the count of occurrences in the cell having Level i in X_1, Level j in X_2, etc.

and $\sum_{ijklmn}\cdots\sum C_{ijklmn} = N_{obs}$ is the total number of observations in the data set being analyzed.

Additional insight into the meaning of H(X) and mutual information and their calculation comes from a different version of Equation 8:

$$H(X) = \sum_{ijklmn}\cdots\sum \frac{C_{ijklmn}}{\sum_{ijklmn}\cdots\sum C_{ijklmn}} \log_2 \left[\frac{\sum_{ijklmn}\cdots\sum C_{ijklmn}}{C_{ijklmn}}\right]$$

$$= \left[\frac{1}{\sum_{ijklmn}\cdots\sum C_{ijklmn}}\right] \sum_{ijklmn}\cdots\sum \left[C_{ijklmn} \log_2\left(\sum_{ijklmn}\cdots\sum C_{ijklmn}\right) - C_{ijklmn} \log_2 C_{ijklmn}\right]$$

$$= \left(\frac{1}{N_{obs}}\right) \sum_{ijklmn}\cdots\sum C_{ijklmn} \log_2 N_{obs} - \left(\frac{1}{N_{obs}}\right) \sum_{ijklmn}\cdots\sum C_{ijklmn} \log_2 C_{ijklmn}$$

$$H(X) = \underbrace{\log_2 N_{obs}}_{\substack{\text{Information}\\ \text{if each}\\ \text{observation}\\ \text{were in a}\\ \text{unique cell}}} - \underbrace{\left(\frac{1}{N_{obs}}\right) \sum_{ijklmn}\cdots\sum C_{ijklmn} \log_2 C_{ijklmn}}_{\substack{\text{Information loss due to clustering}\\ \text{of the observations into a subset}\\ \text{of cells}}} \qquad (10)$$

The entropy of X is expressed in Equation (10) as the difference between two terms. The first, $\log_2 N_{obs}$, is the maximum possible information associated with the given number of observations, i.e., the information that would be present if each observation were unique and occupied a unique cell in the signal space. The second term represents the information that is lost by any clustering of observations into a subset of cells.

Through use of conditional probabilities such as:

$$P_X(ijklmn) = P(X_{1i})P(X_{2j}|X_{1i})P(X_{3k}|X_{1i},X_{2j})\ldots P(X_{6n}|X_{1i},X_{2j}X_{3k},X_{4l},X_{5m})$$

we can have a variety of expressions for H(X):

$$H(X) = H(X_1) + H(X_2|X_1) + H(X_3|X_1,X_2) + \ldots + H(X_6|X_1,X_2,X_3,X_4,X_5) \qquad (11a)$$

$$H(X) = H(X_6) + H(X_2|X_6) + H(X_3|X_2,X_6) + \ldots + H(X_1|X_2,X_3,X_4,X_5,X_6) \qquad (11b)$$

etc.

3.2.3.3 Spectral Band Subsetting. The selection of subsets of spectral bands is a special case of the mutual information expression,

$$I_M(X;Y) = H(X) - H(X|Y)$$

where Y now is a subset X' of the X variables, so

$$I_M(X;X') = H(X) - H(X|X')$$

Whenever a variable, say X_p, is retained, its conditional probability term becomes unity, its contribution to $H(X|X')$ is reduced to zero, and its information content is retained as mutual information. Whenever a variable, say X_q, is eliminated, there is a loss of mutual information. This loss is represented by the conditional entropy term through all conditional probability components in which X_q occurs on the left-hand side of the conditional probability indicator line but not on the right-hand hand (or given) side. The family of entropy relationships illustrated by Equations (11a) and (11b) help define the required calculations.

Single-Band Subsets. The mutual information represented by single-band subsets is:

$$I_M(X;X_p) = H(X_p) = \sum_{\alpha=1}^{N_p} P(X_{p\alpha}) \log_2 \frac{1}{P(X_{p\alpha})} \qquad (12)$$

where p is the band selected and α is the corresponding subscript which indicates the level. This term can be computed from a histogram of signal levels from the band of interest. Alternatively, it can be expressed in terms of the total signal space represented by the data set:

$$I_M(X;X_p) = \sum_{ijklmn}\cdots\sum P_X(ijklmn) \log_2 \frac{1}{P(X_{p\alpha})}$$

where $$P_X(ijklmn) = \frac{C_{ijklmn}}{N_{obs}}$$

and, for example,

$$P(X_{1i}) = \frac{\sum_{jklmn}^{N_2}\cdots\sum^{N_6} C_{ijklmn}}{N_{obs}} \tag{13}$$

Expanding this expression, for X_1 as an example, we have:

$$I_M(X;X_1) = \log_2 N_{obs} - \left(\frac{1}{N_{obs}}\right) \sum_{ijklmn}^{N_1}\cdots\sum^{N_6} C_{ijklmn} \log_2 \left(\sum_{jklmn}^{N_2}\cdots\sum^{N_6} C_{ijklmn} \right) \tag{14a}$$

$$I_M(X;X_1) = \log_2 N_{obs} - \left(\frac{1}{N_{obs}}\right) \sum_{i}^{N_1} \left(\sum_{jklmn}^{N_2}\cdots\sum^{N_6} C_{ijklmn} \right) \log_2 \left(\sum_{jklmn}^{N_2}\cdots\sum^{N_6} C_{ijklmn} \right) \tag{14b}$$

Note the similarity between Equations (14) and (12), the only difference being in the second logarithmic term on the right-hand side of the equation. In Equation (12) this term involves the count in a single cell, while in Equation (14) it is the sum of counts in all cells that have a given X_{1i} level, i.e., the counts are summed over all excluded variables. The pattern holds for all other combinations of variables, as shown next for five-band subsets.

Five-Band Subsets. Choosing a subset of five from six bands is the same as choosing to eliminate the sixth. The conditional entropy in the case of eliminating Band X_1 is:

$$H(X|X_2,X_3,X_4,X_5,X_6) = \sum_{ijklmn}\cdots\sum P_X(ijklmn) \log_2 \frac{1}{P_X(i|jklmn)} \tag{15a}$$

$$= \left(\frac{1}{N_{obs}}\right) \sum_{ijklmn}\cdots\sum C_{ijklmn} \log_2 \frac{\sum_{i=1}^{N_1} C_{ijklmn}}{C_{ijklmn}} \tag{15b}$$

$$= \left(\frac{1}{N_{obs}}\right) \sum_{ijklmn}\cdots\sum C_{ijklmn} \log_2 \left(\sum_{i=1}^{N_1} C_{ijklmn} \right) - \left(\frac{1}{N_{obs}}\right) \sum_{ijklmn}\cdots\sum C_{ijklmn} \log_2 C_{ijklmn} \tag{15c}$$

Since

$$I_M(X;X_2,\ldots,X_6) = H(X) - H(X|X_2,\ldots,X_6)$$

we have, from Equations (10) and (15c),

$$I_M(X;X_2,X_3,X_4,X_5,X_6) = \log_2 N_{obs} - \left(\frac{1}{N_{obs}}\right)\sum_{ijklmn}^{N_1}\cdots\sum^{N_6} C_{ijklmn}\log_2\left(\sum_{i=1}^{N_1} C_{ijklmn}\right) \quad (16a)$$

$$= \log_2 N_{obs} - \left(\frac{1}{N_{obs}}\right)\sum_{jklmn}^{N_2}\cdots\sum^{N_6}\left(\sum_{i=1}^{N_1} C_{ijklmn}\right)\log_2\left(\sum_{i=1}^{N_1} C_{ijklmn}\right) \quad (16b)$$

Again, the form is similar to that of Equations (12) and (14), with the summation in the second logarithmic term being over the excluded variable. The pattern for subsets of two, three, and four bands should now be clear as well.

3.2.4 PRELIMINARY RESULTS AND DISCUSSION

Figure 1 presents information measures for two different quantities, as a function of the number of data variables. First, the system-design capacities of the Landsat-4 TM and MSS are presented, in terms of the number of bits transmitted to the ground and/or recorded on computer-compatible tapes (CCT's). For TM, the number of bits recorded on CCT's is the same as that transmitted (8 bits/channel). For MSS, however, the six-bit telemetered data are expanded to seven bits on the CCT's, with only an apparent gain of information. Nevertheless, most subsequent comparisons involving MSS will use seven-bit data since that is the form in which we have them. Second, the data-space volumes spanned by TM and MSS data from the North Carolina agricultural scene are displayed. These numbers were computed by summing the bit-equivalent of the data-value range (max - min + 1) in each band being considered.

The greater information potential of the TM system design, as compared to the MSS system, is quantified as 48 vs. 24 bits in telemetered data. Upon comparing the fractions of their total data-space volumes that are spanned by data from the agricultural scene, one observes that the TM data fall nine bits short of capacity while the MSS data fall two to six bits short, depending on which curve is used as the reference.

Figure 2 compares the data-space volumes spanned by original and transformed versions of signals from the agricultural scene. The transformations used here are the linear-combination Tasseled-Cap (TASCAP) transformations of MSS [3] and TM [4] data, whose principal variables are Brightness and Greenness. It appears that a bit-rate reduction of about 3 bits/pixel could be achieved for this agricultural scene, without loss of information (See discussions of Figures 3 and 4), by transmitting values from the transformed variables instead of from the original bands.

While the data volumes spanned are quite large, the information-measure values are much smaller, less than 14 bits total (constrained by sample size) for these agricultural data, as shown in Figure 3. This figure compares the agricultural information content of original and TASCAP variables from TM and MSS for the North Carolina scene. In each case, the best subset of each size was used. The mutual information measure which is plotted reflects the actual data-cell patterns into which data from the scene were concentrated and dispersed. For both data sets, relatively little information is gained by the inclusion of more than three variables. The information content of TM data is seen to be from one to more than two bits greater than that of MSS data from the agricultural scene.

Figure 4 illustrates, for the simulated MSS data set, the fact that the information contents of original band values and two types of transformed variables are essentially identical. In addition to TASCAP variables, principal-component variables were extracted and their information content measured. This equality is in keeping with results of theoretical analyses.

Mutual information values for the best and worst band subsets of each size are presented in Figure 5, to illustrate the range of information conveyed by various subsets of the data. The differences are greatest among pairs of variables for both TM and MSS. Figure 6 is a similar comparison for TASCAP variables. In this case, we find an even greater disparity between best and worst, due to the decreased information content of the last TASCAP variables.

The above comparisons have been made primarily using the sets of real TM and MSS data from the agricultural scene. Also analyzed was a simulated data set generated from field-measured reflectance spectra of agricultural crops and soils. Figures 7 and 8 present comparisons of data volumes and information contents of the real and simulated TM and MSS data sets, respectively. Data volumes of the simulated sets are slightly higher and mutual information values slightly lower than for the real data. Figures 9 and 10 present the information ranges spanned by the best and worst subsets for Band and TASCAP variables, respectively. The trends are very similar to those observed for real data (Figures 5 and 6).

Results of analyzing the information-measure values for the various data subsets revealed several interesting trends and produced interesting comparisons with variance-based measures, not all being consistent with the author's initial expectations. When the best subsets were chosen, the information level appeared to be more a function of the number of variables than of the type of variables (i.e., original vs. transformed).

With TM, little difference was found among information measures for the three pairings of TM Brightness, Greenness, and Third-Component, with the TM Brightness/Third-Component vs. TM Greenness/Third-Component comparison being 11.1 vs. 11.5 bits for the real data set, as shown in Figure 11. A greater difference among TM band pairs is evident in Figure 12. While the proportion of variance explained by the first two principal components of MSS data was essentially unity, the information measure showed a lower percentage of the total information was in these two components. The third MSS TASCAP variable (Yellowness) also showed a greater information increment than we have come to expect based on experience in comparing eigenvalues and viewing scatter diagrams of MSS Greenness vs. Yellowness for agricultural data (likely a result of the several-count range of values in the Yellowness variable, i.e., of the thickness of the principal Brightness-Greenness plane). However, the information measure for the MSS Greenness-Yellowness pair was substantially lower than for the MSS Brightness-Greenness pair (7.5 vs. 9.5 bits for the real data set), which is consistent with those prior expectations. The above results indicate a greater data dispersion (and information potential) for the Third Component of TM than of MSS. Also, correlations for TM of -0.69 and 0.36 were noted between Third-Component values and Brightness and Greenness values, respectively, whereas they were uncorrelated for MSS. This is consistent with another examination of this agricultural data set which revealed a somewhat planar TM dispersion pattern that is not aligned with any TM TASCAP axis (although the use of the TASCAP coordinates can still markedly assist interpretation and analysis of the data values).

These results and observations are considered to be preliminary in nature and the reader is urged not to treat them as final, especially since the possibility for data-set dependence exists and only one real and one simulated data set were analyzed here. The information measure employed measures the number of data cells occupied and their populations, independent of their class membership, but consequently is dependent on the population composition of the samples that comprise the data sets. Thus, all results must be interpreted in light of the data populations analyzed. In the simulated data sets, for instance, vegetation samples vastly outnumbered bare-soil samples. It is noted that the measure also is independent of the noise levels in the various bands and of the ease and consistency of interpretation of the spectral variables (an advantage ascribed to TASCAP variables), which are other factors which should be considered. The presence of noise adds variance and could make the apparent information content greater than the true information content of ideal signals.

3.3 SIGNIFICANT RESULTS

An information-theoretic measure of the information content of various subsets of multispectral variables was developed and applied to a real and a simulated set of simultaneous TM and MSS data from agricultural crops and soils. Preliminary observations and comparisons are made.

3.4 PUBLICATIONS AND PRESENTATIONS

A paper describing results of this investigation was invited for presentation and publication at the 1984 Purdue/LARS Symposium on Machine Processing of Remotely Sensed Data, June 12-14, 1984. Entitled "Thematic Mapper Radiometric Characterization", and co-authored by William A. Malila and Michael D. Metzler, it will be presented in a session on TM Data Quality Analysis which is to be chaired by W. Malila.

3.5 RECOMMENDATIONS

None.

3.6 FUNDS EXPENDED

A total of approximately   was expended during the three months November 1983 through February 1984. The cumulative spending through February represents approximately 62% of the total contract award. Expenditures during the period 1-20 March 1984 are not included in this percentage value.

REFERENCES

1. Malila, W.A., M.D. Metzler and E.P. Crist. Study on Spectral/Radiometric Characteristics of the Thematic Mapper for Land Use Applications, Fourth Type II Quarterly Status and Technical Report, NASA/GSFC, ERIM Report 164000-4-T, ERIM, Ann Arbor, MI, 1983.

2. Crist, E.P. "Comparison of Coincident Landsat-4 MSS and TM Data Over an Agricultural Region", Technical Papers of the 50th Annual Meeting of the American Society of Photogrammetry, Vol. 2, pp. 508-517, 1984.

3. Kauth, R.J. and G.S. Thomas, "The Tasseled Cap -- A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat", Symposium on Machine Processing of Remotely Sensed Data, June 29 - July 1, 1976, Purdue University, LARS, W. Lafayette, IN, 1976.

4. Crist, E.P. and R.C. Cicone, "A Physically-based Transformation of Thematic Mapper Data - The TM Tasseled Cap", IEEE Transactions on Geoscience and Remote Sensing, May 1984.

FIG. 1 COMPARISON OF TM AND MSS INFORMATION CAPACITIES

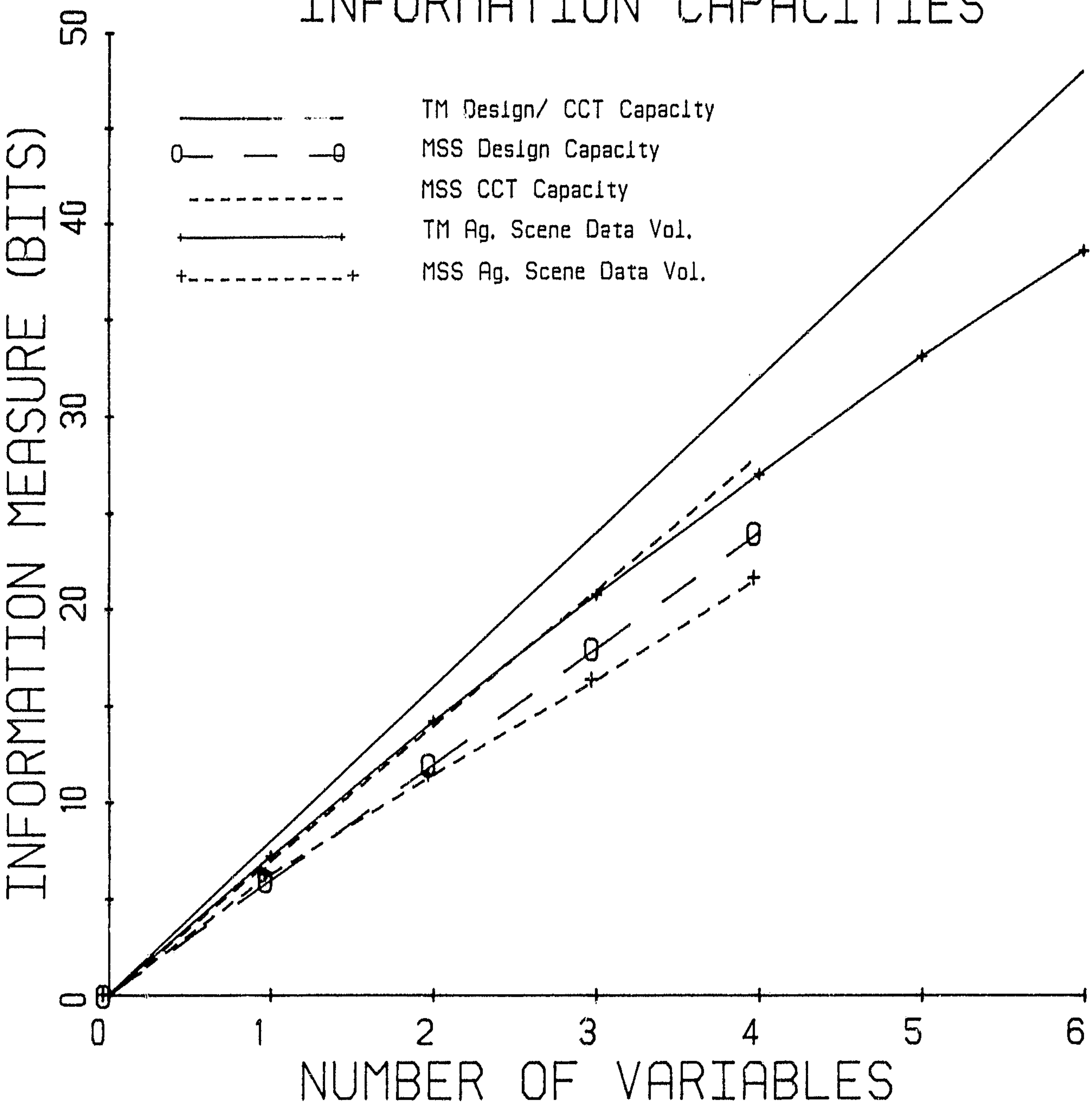
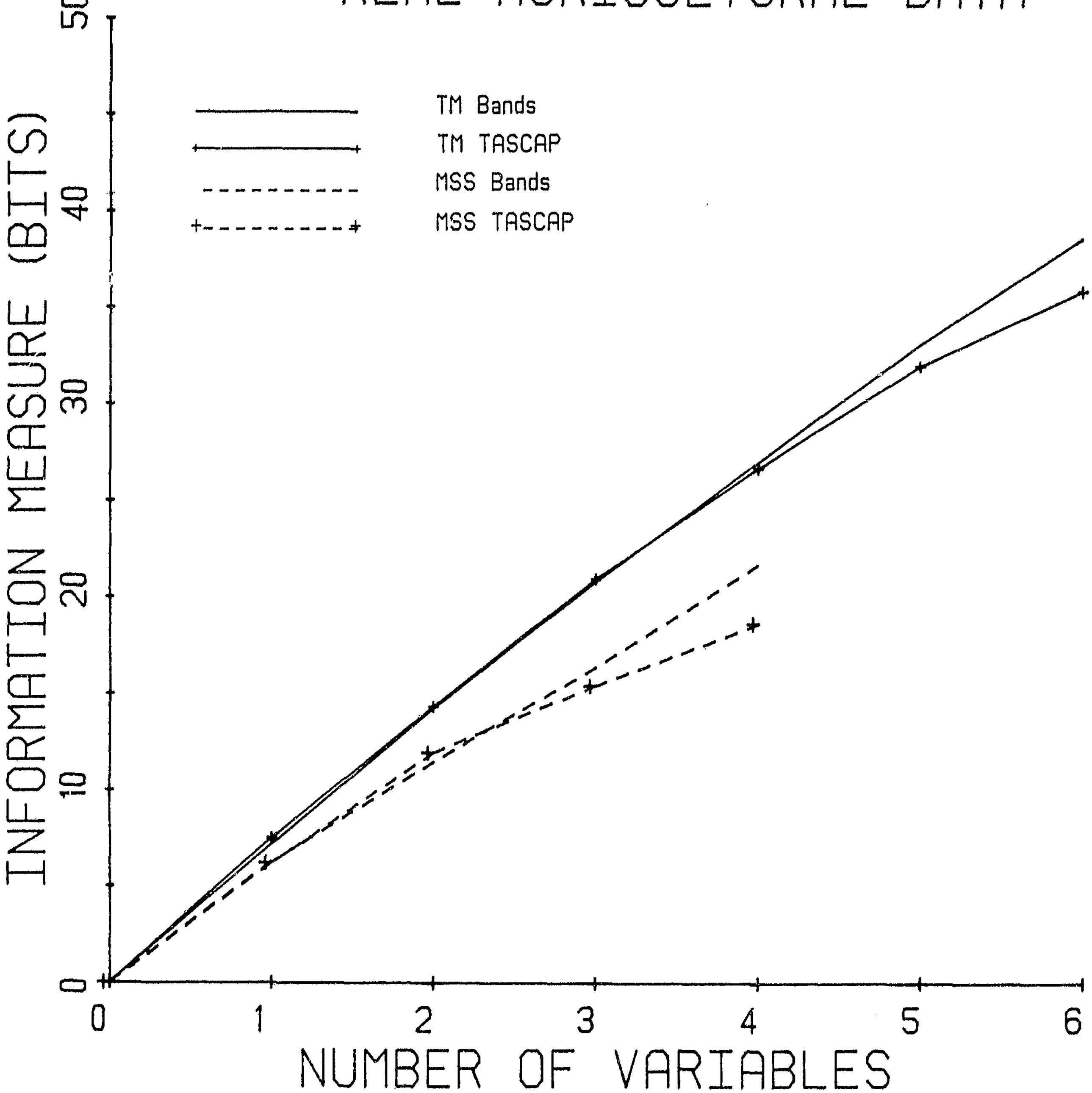


FIG. 2 COMPARISON OF DATA VOLUMES:
REAL AGRICULTURAL DATA



ORIGINAL PAGE IS
OF POOR QUALITY

FIG. 3 COMPARISON OF AGRICULTURAL INFORMATION CONTENT REAL DATA

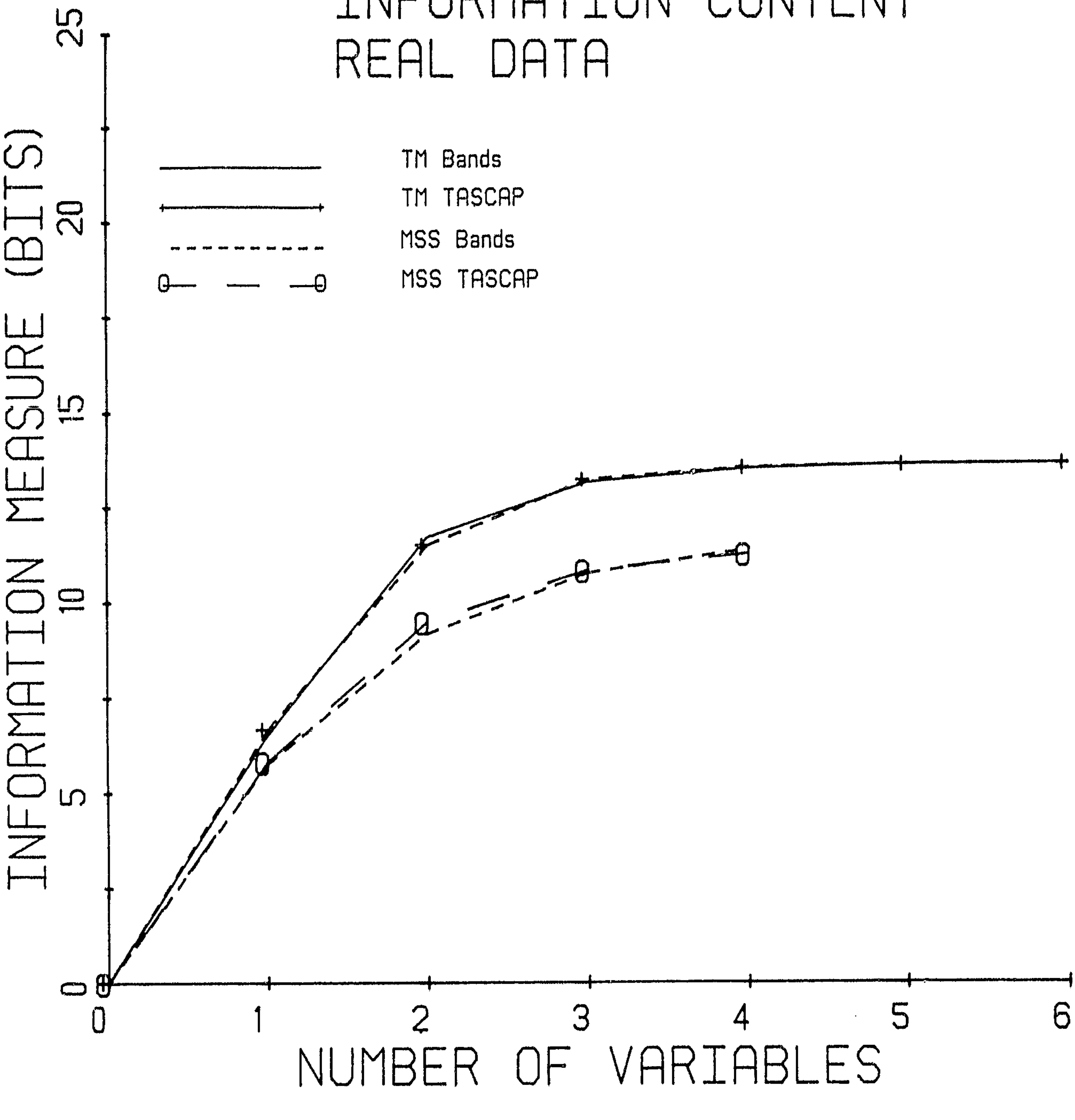


FIG. 4 INFORMATION COMPARISON OF MSS BAND, TASCAP, AND PRINCIPAL-COMPONENT VARIABLES (SIMULATED AG. DATA)

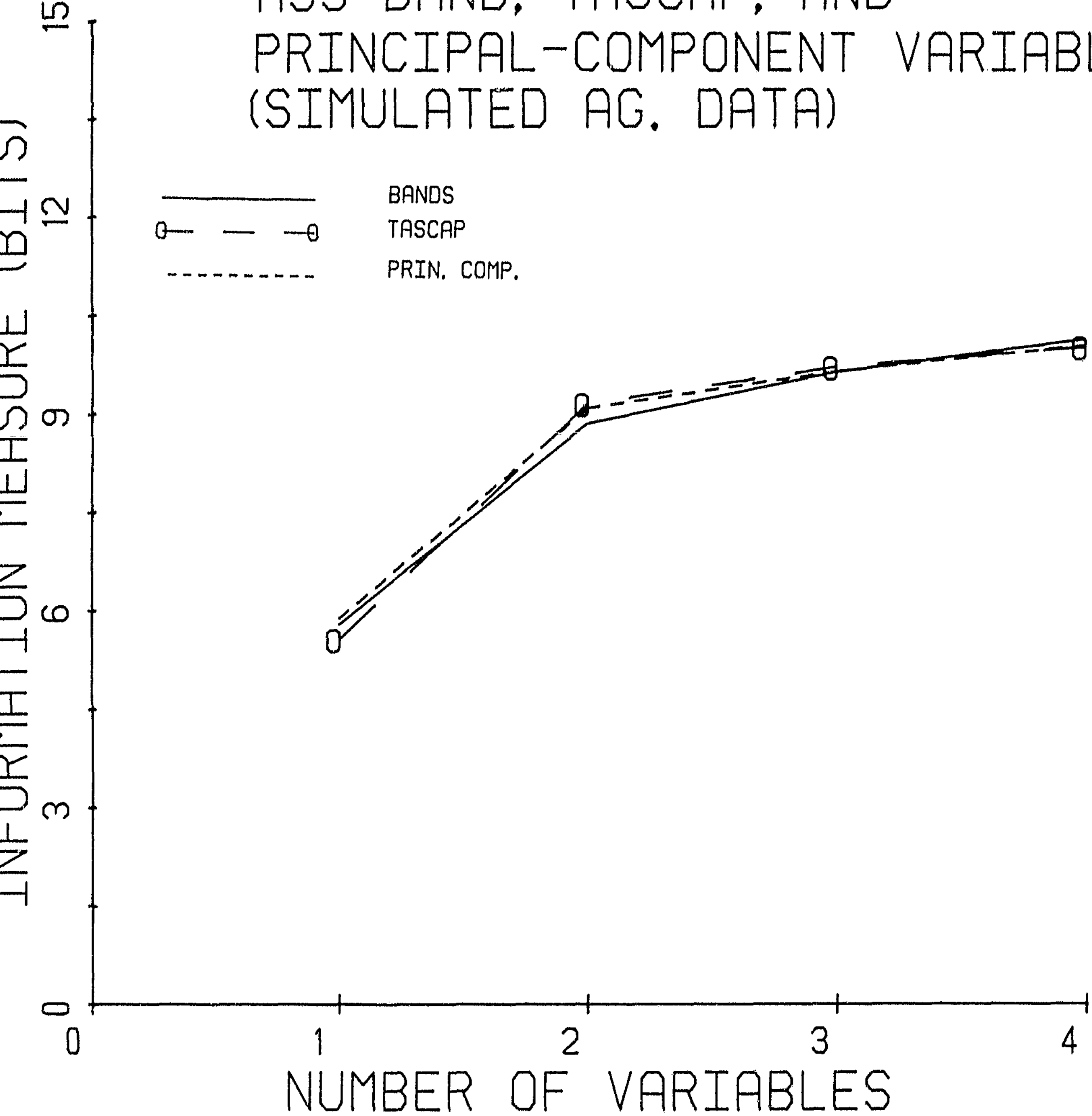
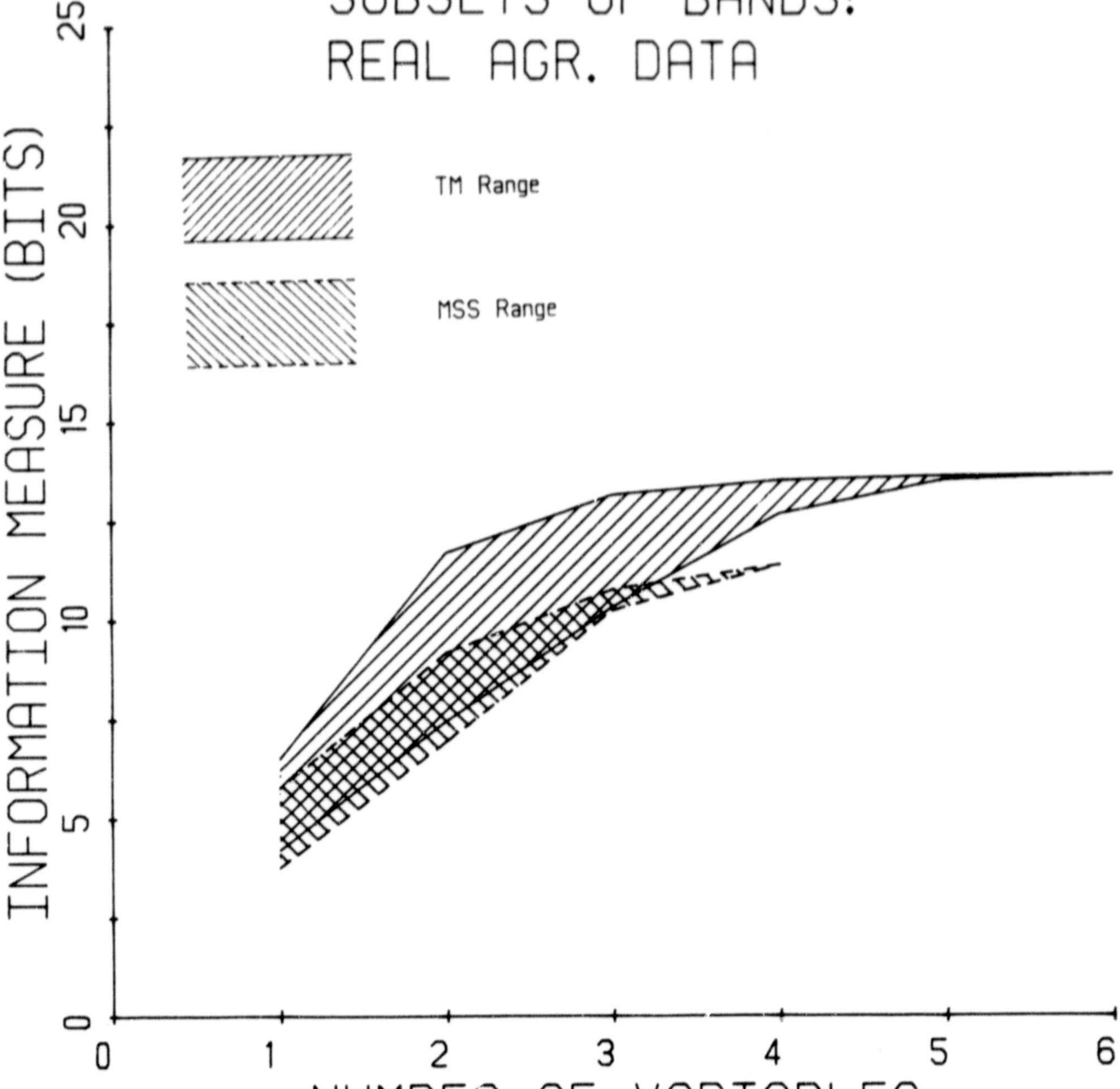


FIG. 5 RANGE OF INFORMATION IN SUBSETS OF BANDS: REAL AGR. DATA



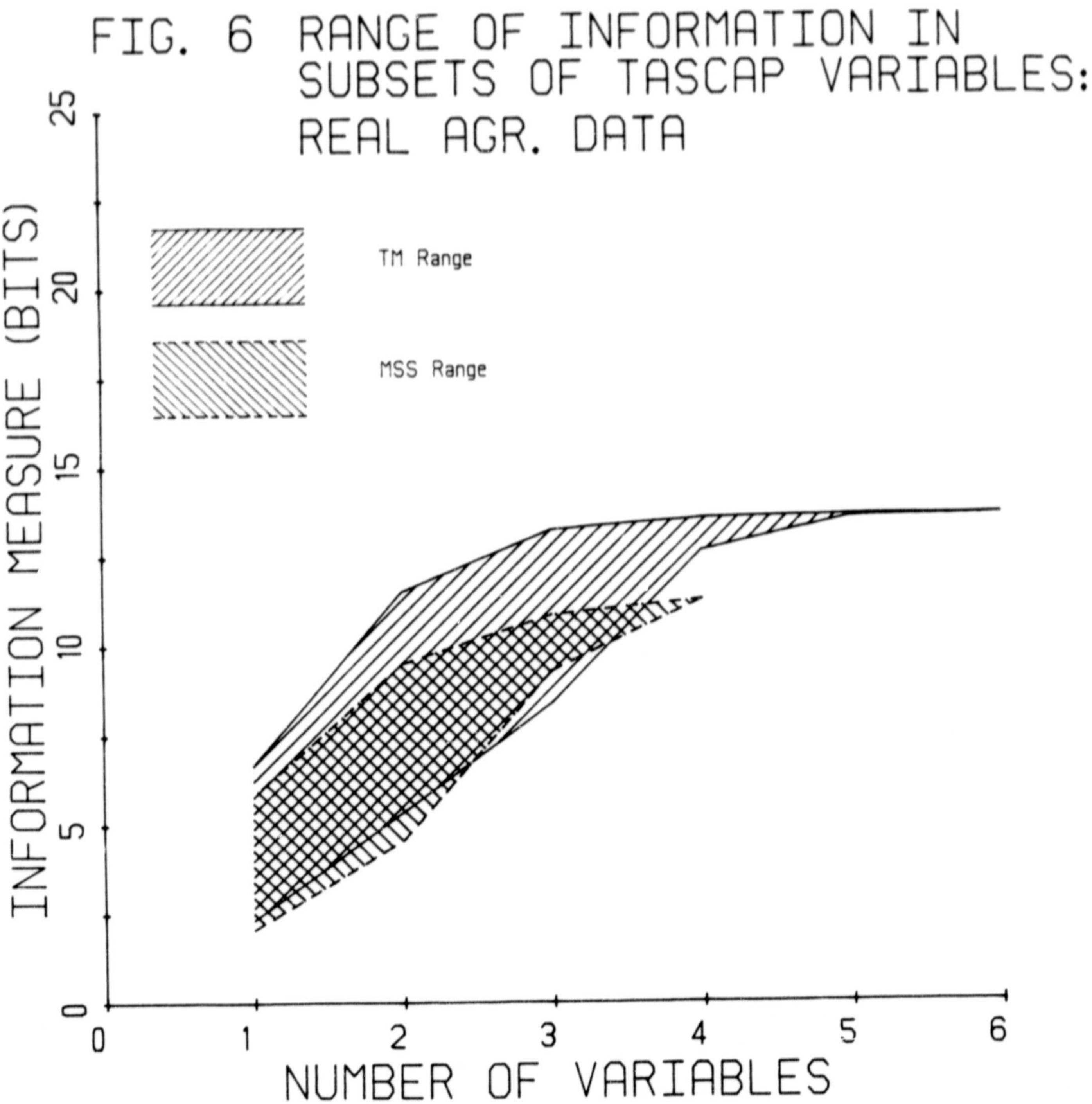


FIG. 6 RANGE OF INFORMATION IN SUBSETS OF TASCAP VARIABLES: REAL AGR. DATA
INFORMATION MEASURE (BITS)
0
5
10
15
20
25
TM Range
MSS Range
0
1
2
3
4
5
6
NUMBER OF VARIABLES

FIG. 7 COMPARISON OF REAL AND SIMULATED AGRICULTURAL DATA SET: THEMATIC MAPPER

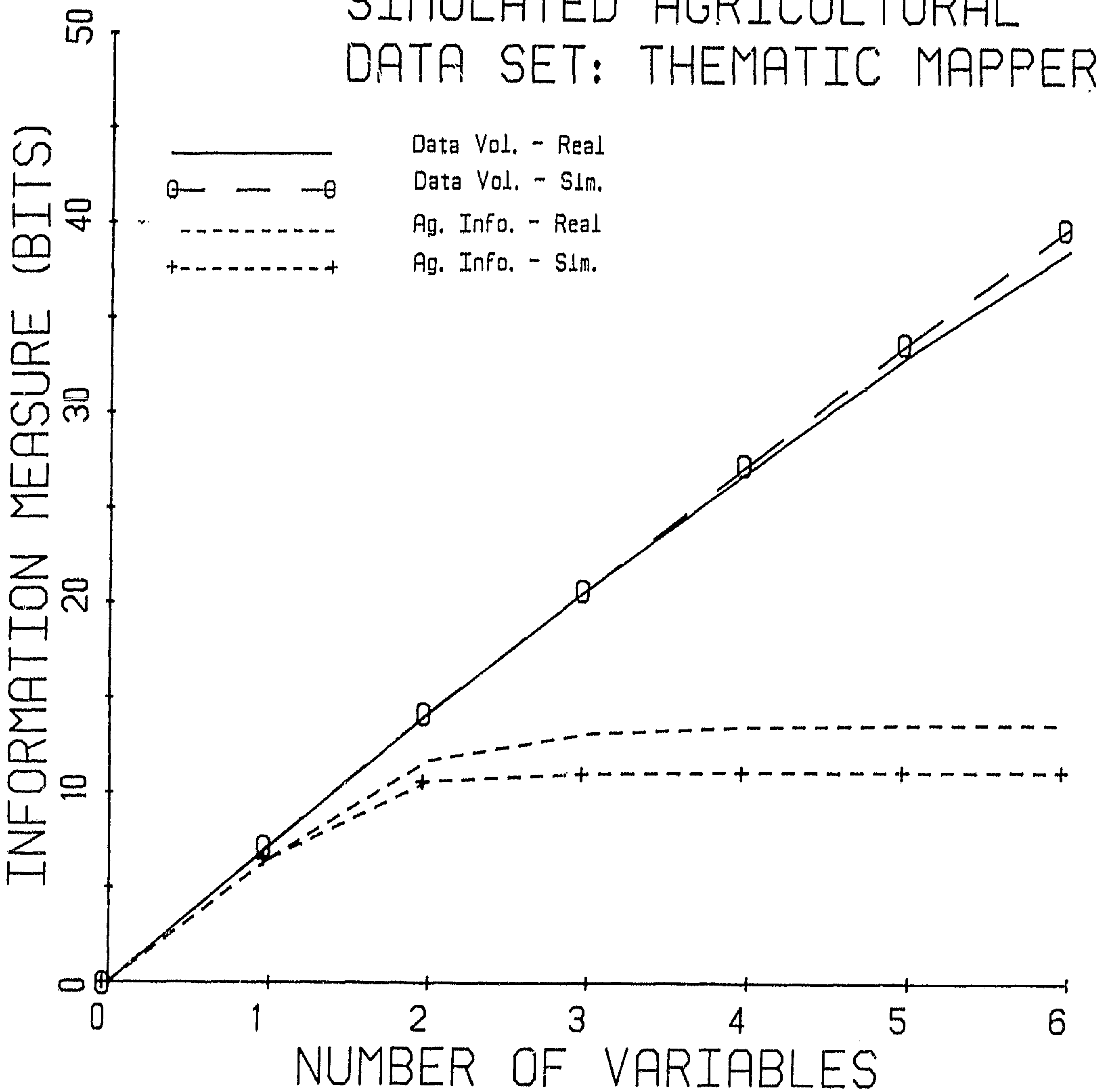
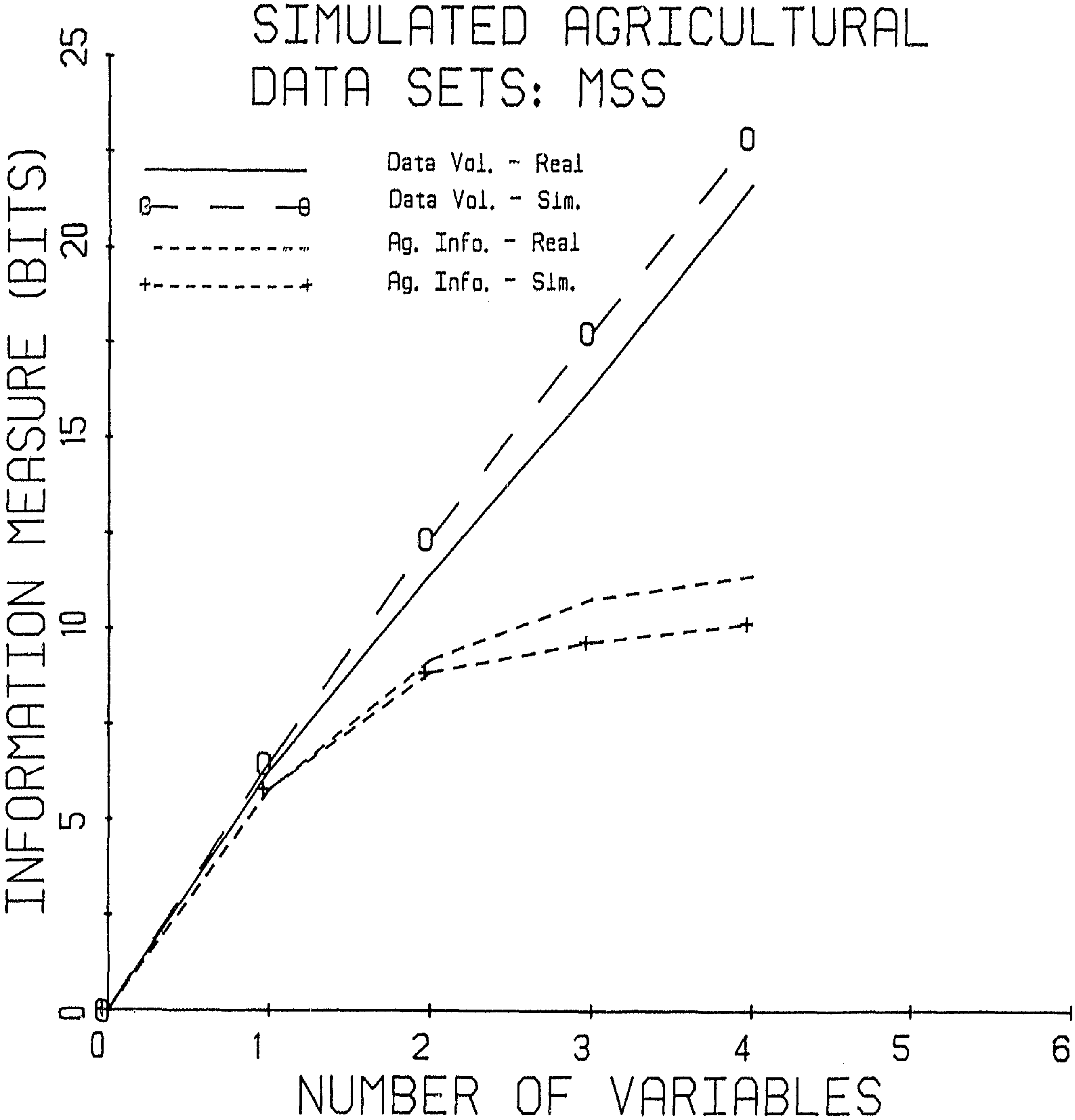


FIG. 8 COMPARISON OF REAL AND SIMULATED AGRICULTURAL DATA SETS: MSS



ORIGINAL PAGE IS
OF POOR QUALITY

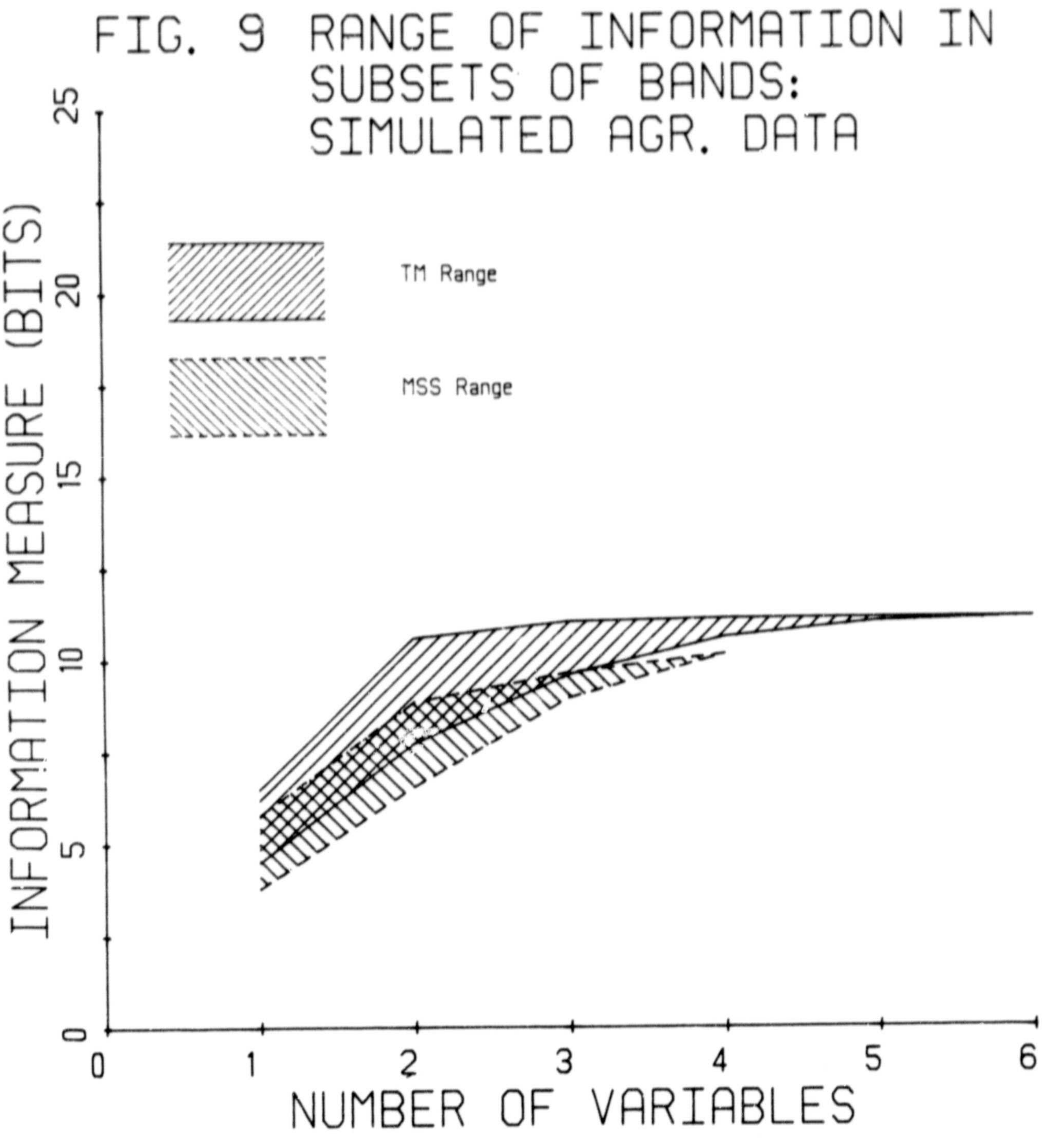
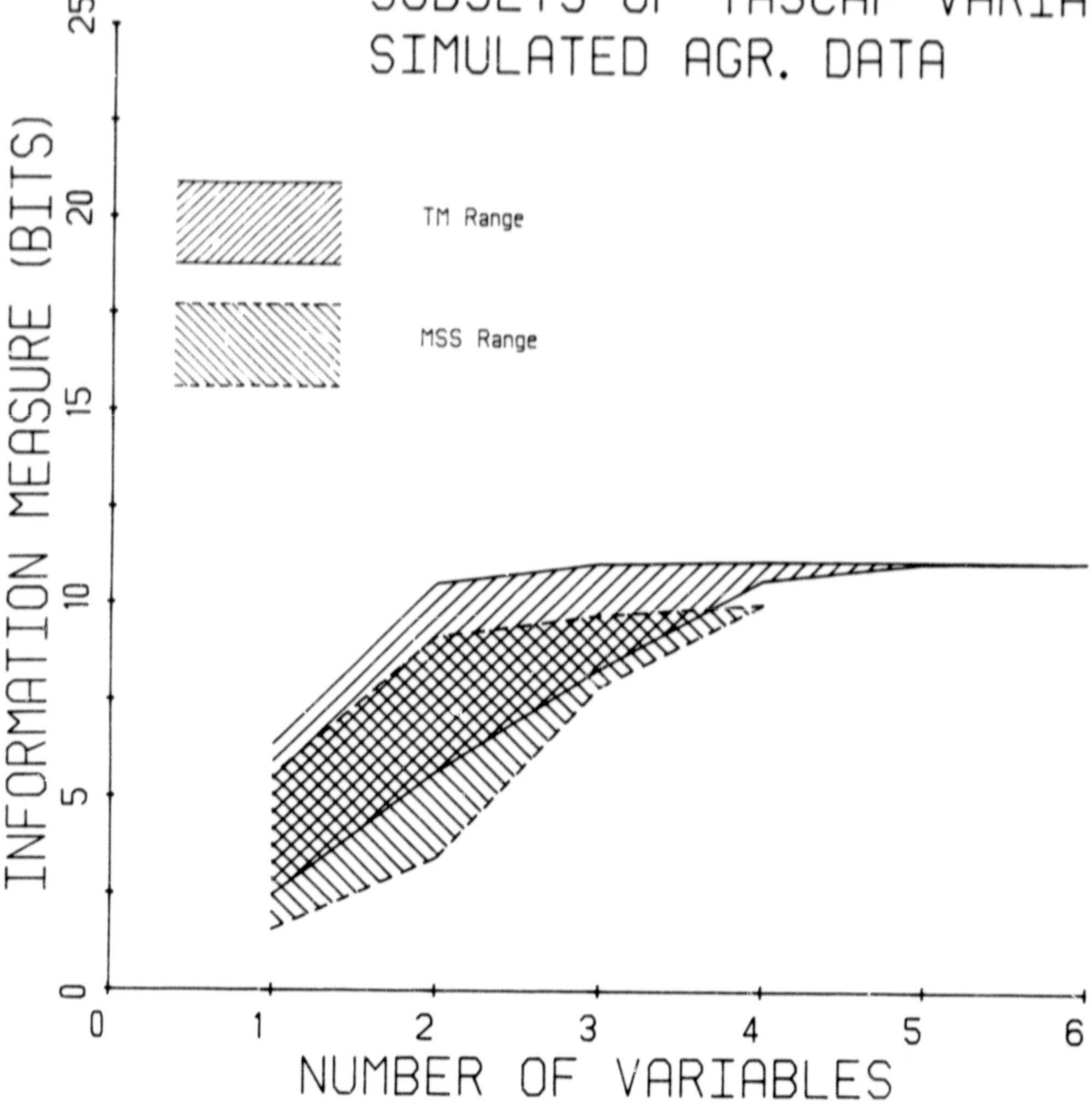


FIG. 10 RANGE OF INFORMATION IN SUBSETS OF TASCAP VARIABLES SIMULATED AGR. DATA



ORIGINAL PAGE IS
OF POOR QUALITY

FIG. 11 TM-MSS TASCAP COMPARISONS:
REAL AGRICULTURAL DATA

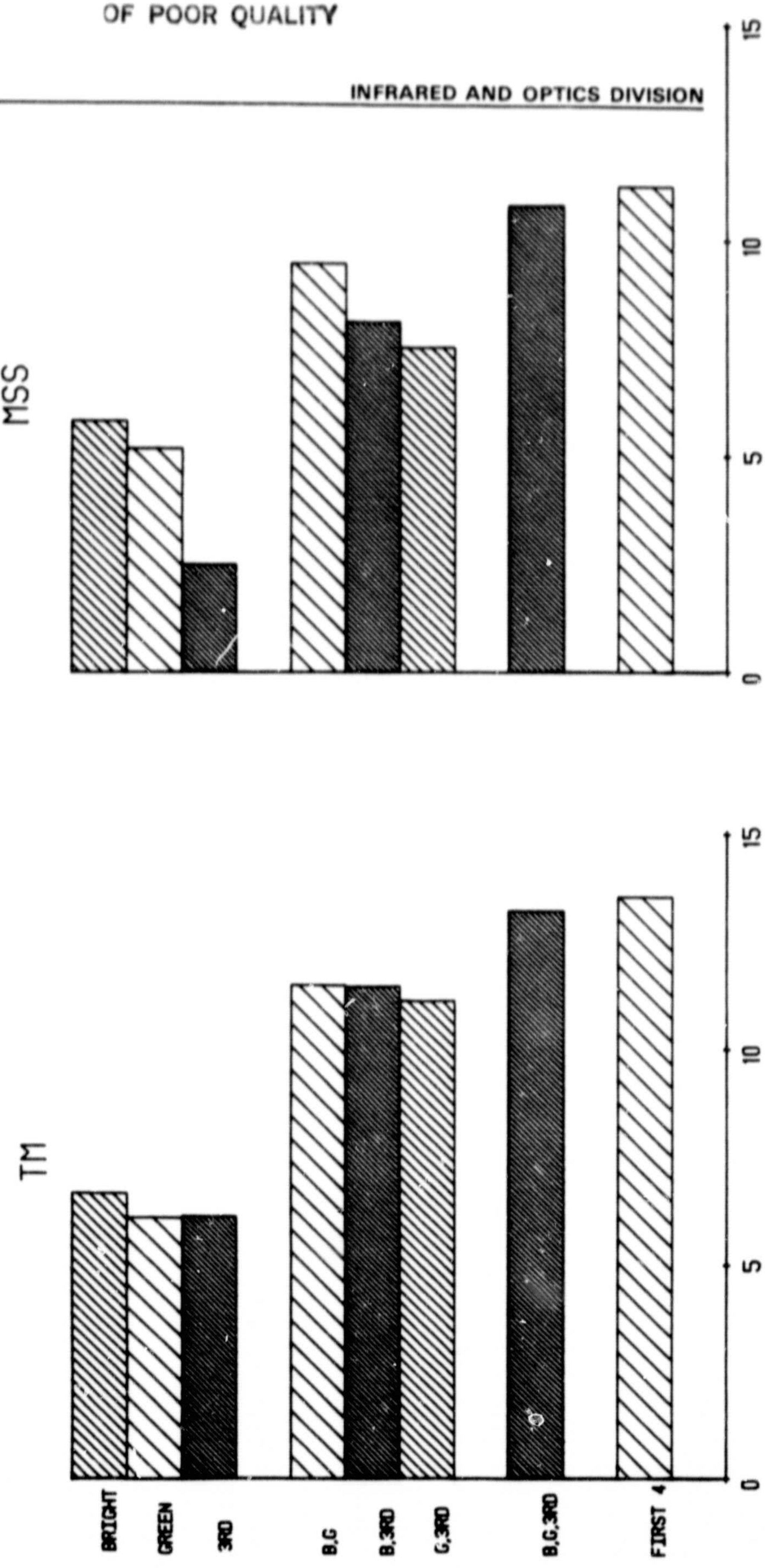


FIG. 12 TM-MSS BAND COMPARISONS: REAL AGRICULTURAL DATA